Cauchy-Schwarz par le calcul différentiel
Dominique Hoareau, 26-10-2003
domeh@wanadoo.fr

On propose de démontrer le théorème du multiplicateur de Lagrange (ou théorème des extrema liés), accessible dès que l´on dispose du théorème des fonctions implicites dans $\mathbb{R}^2$, puis on l´utilise pour justifier l´inégalité de Cauchy-Schwarz.

Fonctions implicites : cas de deux variables réelles
Enoncé : Soit U un ouvert non vide de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ une fonction C^1, $(a,b) \in U$ tel que : $f(a,b) = 0$. On suppose que $\frac{\partial f}{\partial y}(a,b) \neq 0$.
Alors il existe :

- des voisinages ouverts I et J de a et b
- une fonction $\varphi : I \to \mathbb{R}$ C^1

tels que : $\forall (x,y) \in U \quad [(x,y) \in I \times J \; et \; f(x,y) = 0] \Leftrightarrow [x \in I \; et \; y = \varphi(x)]$.
La dérivation licite par rapport à x dans la relation $f(x, \varphi(x)) = 0$ donne :

$$\forall x \in I \quad \varphi'(x) = \frac{-\frac{\partial f}{\partial x}(x,\varphi(x))}{\frac{\partial f}{\partial y}(x,\varphi(x))}.$$

Commentaire : Sous les conditions du théorème, après un zoom sur le point (a,b), la ligne de niveau 0 de f est, dans la "lucarne" $I \times J$ (localement autour de (a,b)), le graphe d´une fonction φ de classe C^1. Sa tangente au point a est dirigée par le vecteur $\left(1, \frac{-\frac{\partial f}{\partial x}(a,b)}{\frac{\partial f}{\partial y}(a,b)}\right)$ qui est orthogonal à la direction $(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b)$. La condition $\frac{\partial f}{\partial y}(a,b) \neq 0$ s´interprète donc intuitivement comme la donnée d´une "tangente non verticale" en (a,b) pour la courbe d´équation $f(x,y) = 0$.

1 Théorème du multiplicateur de Lagrange

Énoncé
Soit U un ouvert non vide de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ deux fonctions C^1. On pose : $M = \{x \in U \; ; \; g(x) = 0\}$. Soit $a \in M$ tel que $dg_a \neq 0$. Si $f_{|M}$ présente en a un extrémum local, alors il existe un réel ν, appelé multiplicateur de Lagrange, tel que $df_a = \nu dg_a$.

Interprétation
En tout point a extrémum de f sur M, la ligne de niveau $f(a)$ de f et M ont même "espace tangent en a " ($ker \; df_a = ker \; dg_a$). En notant $< \, , >$ le

produit scalaire usuel de $\mathbb{R}^n$, on peut dire que les vecteurs gradients $\bigtriangledown f_a$ et $\bigtriangledown g_a$ définis par $df_a =< \bigtriangledown f_a, \bullet >$ et $dg_a =< \bigtriangledown g_a, \bullet >$, sont colinéaires.

Preuve
On choisit v dans $\mathbb{R}^n$ tel que : $dg_a.v \neq 0$.
Soit $u \in \mathbb{R}^n$. On pose :

$$\Omega = \{(\lambda, \mu) \in \mathbb{R}^2\,;\, a + \lambda u + \mu v \in U\}.$$

Puisque U est un ouvert de $\mathbb{R}^n$ contenant a et $(\lambda, \mu) \mapsto a + \lambda u + \mu v$ est continue sur $\mathbb{R}^2$, Ω est un ouvert de $\mathbb{R}^2$ contenant $(0,0)$. Soit Ψ l´application $(\lambda, \mu) \mapsto g(a + \lambda u + \mu v)$ de Ω dans $\mathbb{R}$. Ψ est C^1 sur Ω car composée de fonctions C^1. De plus : $\frac{\partial \Psi}{\partial \mu}(0,0) = dg_a.v \neq 0$ donc le théorème des fonctions implicites permet d´exprimer localement μ en fonction de λ : on choisit $\alpha > 0$ et $h :]-\alpha, \alpha[\to \mathbb{R}$ C^1 tels que :

$$\forall \lambda \in]-\alpha, \alpha[\ (\lambda, h(\lambda)) \in \Omega; h(0) = 0; \Psi(\lambda, h(\lambda)) = 0.$$

On considère alors l´application $F : \lambda \mapsto f(a+\lambda u+h(\lambda)v)$ de $]-\alpha, \alpha[$ dans $\mathbb{R}$. On a déplacé le problème initial vers un problème d´une seule variable. F est C^1 et présente un extrémum local en 0 (intérieur à $]-\alpha, \alpha[$) donc : $F'(0) = 0$. Par ailleurs : $F'(0) = df_a.(u+h'(0)v) = df_a.(u - \frac{\frac{\partial \Psi}{\partial \lambda}(0,0)}{\frac{\partial \Psi}{\partial \mu}(0,0)}v) = df_a.(u - \frac{dg_a.u}{dg_a.v}v)$ donc $df_a.u = df_a.(\frac{dg_a.u}{dg_a.v}v) = \frac{df_a.v}{dg_a.v}dg_a.u$. On a trouvé $\nu = \frac{df_a.v}{dg_a.v}$ indépendant de u tel que : $df_a.u = \nu dg_a.u$. D´où : $df_a = \nu dg_a$.

Voici deux applications du théorème du multiplicateur de Lagrange.

Caractérisation de $SO(n, \mathbb{R})$ dans $SL(n, \mathbb{R})$
(cf *Thèmes de géométrie*, M. Alessandri, Dunod)

$\mathcal{M}_n(\mathbb{R})$ est muni du produit scalaire classique : $< A, B >= trace({}^tAB)$. On veut calculer la distance euclidienne entre la matrice nulle et le fermé $SL(n, \mathbb{R})$ de $\mathcal{M}_n(\mathbb{R})$. Il s´agit donc de minimiser $f : M \mapsto < M, M >$ sous la contrainte $g(M) = det(M) - 1 = 0$.

Soit d cette distance. Pour tout $p \in \mathbb{N}^*$, la boule fermée B_p de centre 0 et de rayon $d + \frac{1}{p}$ rencontre $SL(n, \mathbb{R})$ d´où l´existence d´une suite $(M_p)_{p \geqslant 1}$ de $SL(n, \mathbb{R})$ telle que : $\forall p \in \mathbb{N}^* \quad d \leqslant \| M_p \| \leqslant d + \frac{1}{p}$. Par compacité des B_p (fermés bornés en dimension finie), quitte à en extraire une sous-suite, on peut supposer que (M_p) converge. Sa limite notée M_0 appartient au fermé $SL(n, \mathbb{R}) = det^{-1}(\{1\})$ et vérifie $< M_0, M_0 >= d^2$ (passage à la limite). On retiendra qu´en dimension finie, la distance à un fermé est toujours atteinte (argument de locale compacité).

On a de façon classique :

$$df_{M_0} =< 2M_0, \bullet > \quad ; \quad dg_{M_0} =< com(M_0), \bullet >$$

où $com(M_0)$ désigne la comatrice de M_0 (cf infra). Par le théorème du multiplicateur de Lagrange, il existe $\nu \in \mathbb{R}$ tel que : $M_0 = \nu com(M_0)$. On a : $M_0^{-1} = {}^t(com\, M_0)$ donc ${}^tM_0M_0 = \nu I$. Il vient : $det({}^tM_0M_0) = \nu^n = 1$ avec, puisque ${}^tM_0M_0$ symétrique définie positive, $\nu \in \mathbb{R}^{*+}$. De là : $\nu = 1$ et $M_0 \in SO(n, \mathbb{R})$. Or, pour $M \in SO(n, \mathbb{R})$, $< M, M >= n$ donc $d = \sqrt{n}$ et on a caractérisé $SO(n, \mathbb{R})$ dans $SL(n, \mathbb{R})$ comme l´ensemble des éléments de norme euclidienne minimum.

Justification rapide de : $\forall M \in GL_n(\mathbb{R}) \quad d\, det_M =< com\, M, \bullet >$.
Pour $H \in \mathcal{M}_n(\mathbb{R})$, $det(M+H) = det(M)\, det(I+M^{-1}H)$. Or, si on note $E_{i,j}$ la matrice élémentaire $n \times n$ avec des zéros partout sauf un 1 en position (i, j), pour $H = [h_{i,j}]$, on a : $d\, det_I.H = \sum_{1 \leqslant i,j \leqslant n} h_{i,j} d\, det_I.E_{i,j}$. On remarque que : $det(I + E_{i,j}) = det(I)$ si $i \neq j$, $det(I + E_{i,j}) = det(I) + 1$ sinon. D´où : $d\, det_I.H = trace(H)$, et $d\, det_M.H = det(M)\, trace(M^{-1}H)$
$= trace({}^tcom(M)\, H)$. cqfd.

Inégalité arithmético-géométrique
(*les maths en tête*, Gourdon, Ellipses)

Montrer que : $\forall (x_1, ..., x_n) \in (\mathbb{R}^{*+})^n \quad (\prod_{1 \leqslant i \leqslant n} x_i)^{\frac{1}{n}} \leqslant \frac{\sum_{i=1}^n x_i}{n}$.

On envisage $f : (x_1, ..., x_n) \in \mathbb{R}^n \mapsto \prod_{1 \leqslant i \leqslant n} x_i$ et, pour $s > 0$,
$g_s : (x_1, ..., x_n) \in \mathbb{R}^n \mapsto \sum_{i=1}^n x_i - s$. On pose $M_s = g_s^{-1}\{0\}$. f est continue sur le compact $K_s = M_s \cap (\mathbb{R}^+)^n$ (fermé borné de $\mathbb{R}^n$), donc $f_{|K_s}$ admet un maximum *global* atteint en un $a = (a_1, ..., a_n)$ de K_s. a est nécessairement dans l´ouvert relatif $M_s \cap (\mathbb{R}^{*+})^n$ de M_s. a est donc un maximum $relatif$ de f sous la liaison M_s, et par le théorème des extrema liés, il existe $\nu \in \mathbb{R}$ tel que : $\bigtriangledown f_a = \nu. \bigtriangledown (g_s)_a$. On a donc : $\forall i \in \{1, ..., n\} \quad \frac{f(a)}{a_i} = \nu$, ce qui prouve, puisque $f(a) \neq 0$, que les a_i sont égaux : $\forall 1 \leqslant i \leqslant n \quad a_i = \frac{s}{n}$.
Bilan : $\forall (x_1, ..., x_n) \in (\mathbb{R}^{*+})^n \quad \prod_{1 \leqslant i \leqslant n} x_i \leqslant \left(\frac{\sum_{i=1}^n x_i}{n}\right)^n$.

2 Multiplicateur de Lagrange contre Cauchy-Schwarz dans la Diagonalisation des endomorphismes symétriques de $\mathbb{R}^n$

(cf RMS – *décembre* – 1989, J.B Hirriart-Urruty et G. Lion, Vuibert)
On rappelle la version générale de :

Inégalité de Cauchy-Schwarz
Soit B une forme bilinéaire symétrique sur $\mathbb{R}^n$. On suppose que la forme quadratique Q associée à B est positive ($\forall x \in \mathbb{R}^n \quad Q(x) \geqslant 0$).
Pour tous x et y de $\mathbb{R}^n$, on a : $B(x,y)^2 \leqslant Q(x)Q(y)$.

Lorsqu´on étudie la réduction d´un endomorphisme symétrique u de $\mathbb{R}^n$, on commence par prouver que u possède au moins une valeur propre réelle. On note $<,>$ le produit scalaire usuel de $\mathbb{R}^n$ et $\| \quad \|$ la norme associée. On considère la forme bilinéaire symétrique B de $\mathbb{R}^n$ définie par $B(x,y) =< x, u(y) >$, et la forme quadratique Q associée. Q est continue sur le compact $S = \{x \in \mathbb{R}^n ; < x, x >= 1\}$ donc atteint ses bornes sur S. On pose $\lambda = \max_{x\in S} Q(x)$ et on choisit x_0 sur S tel que : $Q(x_0) = \lambda$. La forme quadratique $Q_1 : x \mapsto \lambda \parallel x \parallel^2 - Q(x)$ est positive donc par Cauchy-Schwarz, $B_1(x,x_0)^2 \leqslant Q_1(x)Q_1(x_0)$ où B_1 désigne la forme polaire de Q_1. Avec $Q_1(x_0) = 0$, $\forall x \in \mathbb{R}^n \quad < x, \lambda x_0 - u(x_0) > \leqslant 0$. En particulier pour $x = \lambda x_0 - u(x_0)$, on obtient : $\lambda x_0 - u(x_0) = 0$.

Autre vision de l´affaire (cf *Calcul differentiel*, Avez, Masson) :
x_0 maximise Q sous la contrainte $h(x) =< x, x > -1 = 0$. On a donc $\bigtriangledown Q_{x_0}$ colinéaire à $\bigtriangledown h_{x_0}$, ce qui donne avec u symétrique : $u(x_0)$ colinéaire à x_0.

3 Cauchy-Schwarz et les extrema liés

Version démontrée
$<,>$ désigne un produit scalaire sur $\mathbb{R}^n$ et $\| \quad \|$ la norme associée.
(CS) : pour tous x et y de $\mathbb{R}^n$, $< x, y >^2 \leqslant \parallel x \parallel^2 \parallel y \parallel^2$.

Preuve
Soit $x \in \mathbb{R}^n$.
• si $x = 0$, (CS) est clairement vérifiée pour tout y de $\mathbb{R}^n$.

• On suppose à présent $x \neq 0$. On pose : $S = \{x \in \mathbb{R}^n ; \parallel x \parallel = 1\}$.

↪ On envisage les applications $f : y \mapsto < x, y >^2 - \parallel x \parallel^2 \parallel y \parallel^2$ et $g : y \mapsto \parallel y \parallel^2 - 1$ de $\mathbb{R}^n$ dans $\mathbb{R}$. L´application $< x, \bullet >$, forme linéaire sur $\mathbb{R}^n$, et $y \mapsto \parallel y \parallel^2$ sont continues sur $\mathbb{R}^n$ donc f l´est aussi ; sur le compact S (fermé borné en dimension finie), f atteint sa borne supérieure, par exemple en y_0.

↪ Pour $y \in \mathbb{R}^n$ et $h \in \mathbb{R}^n$,
$f(y+h) = f(y) + \underbrace{2\left\langle < x, y > x - \parallel x \parallel^2 y, h \right\rangle}_{lin\acute{e}aire\, en\, h} + \underbrace{< x, h >^2 - \parallel x \parallel^2 \parallel h \parallel^2}_{N(h)}$.

On remarque : $\frac{N(h)}{\|h\|} = \left\langle x, \frac{h}{\|h\|^{\frac{1}{2}}} \right\rangle^2 - \parallel x \parallel^2 \parallel h \parallel$. Par continuité des

applications $h \mapsto \| h \|$ et $h \mapsto < x, h >^2$ en 0, on a : $\lim_{h \to 0} \frac{N(h)}{\|h\|} = 0$. De là, f est différentiable en y et :

$$df_y = 2 \left\langle < x, y > x - \| x \|^2 y, \bullet \right\rangle.$$

$y \mapsto df_y$ est clairement linéaire de $\mathbb{R}^n$ dans son dual (algébrique, et topologique) donc f est C^1 sur $\mathbb{R}^n$.

On vérifie que g est différentiable en tout y de $\mathbb{R}^n$, $dg_y = < 2y, \bullet >$, et g est C^1 sur $\mathbb{R}^n$.

$\hookrightarrow$ Par le théorème des extréma liés, on trouve $\lambda \in \mathbb{R}$ tel que :

$$2[< x, y_0 > x - \| x \|^2 y_0] = \lambda 2 y_0$$

c-à-d :

$$[\lambda + \| x \|^2] y_0 = < x, y_0 > x.$$

Si $< x, y_0 > = 0$, $f(y_0) = - \| x \|^2 \| y_0 \|^2 \leqslant 0$, et pour tout y de S, $f(y) \leqslant 0$.
Si $< x, y_0 > \neq 0$, on remplace x par $\frac{[\lambda + \|x\|^2]}{< x, y_0 >} y_0$ dans $f(y_0)$ et on trouve : $f(y_0) = 0$, ce qui donne encore :

$$\forall y \in S \quad f(y) \leqslant 0.$$

$\hookrightarrow$ L´inégalité $f(y) \leqslant 0$ vraie sur S l´est encore sur $\mathbb{R}^n \backslash \{0\}$ par homogénéité.
Pour $y \neq 0$, on a : $f\left(\frac{y}{\|y\|}\right) \leqslant 0$, c-à-d : $\frac{1}{\|y\|^2} < x, y >^2 - \| x \|^2 \leqslant 0$, ou encore :

$$< x, y >^2 - \| x \|^2 \| y \|^2 \leqslant 0.$$

Enfin, $f(0) = 0$, ce qui termine la preuve.

Remerciements à Dany-Jack Mercier pour sa lecture attentive du texte.